

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:13:01 ; Search time 32 seconds

(without alignments)
766.238 Million cell updates/sec

Title: US-10-015-967-2

Perfect score: 644
Sequence: 1 MKVLISLLPLPLMLSMV.....SRACQFLKOCOLRSFALPL 119

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPREMBL_21:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriophage:*
17: sp-archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	322	50.0	128	11	08R3U6
2	77.5	12.0	477	5	062022
3	76.5	11.9	411	13	09W745
4	72.5	11.3	3313	11	088278
5	71.5	11.1	108	10	082328
6	71	11.0	188	4	08TEV2
7	71	11.0	3190	5	001368
8	71	11.0	3275	5	09W321
9	70.5	10.9	160	10	09SQH1
10	70.5	10.9	1290	3	0960A3
11	70	10.9	117	4	09BRN8
12	70	10.9	743	5	097120
13	70	10.9	1095	5	021852
14	70	10.9	1664	5	09TV02
15	69.5	10.8	221	12	085299
16	69.5	10.8	468	4	090C32

17	69	10.7	111	10	09FWV9
18	69	10.7	200	16	09HVU8
19	69	10.7	271	4	09HBS0
20	69	10.7	378	3	074213
21	69	10.7	413	10	09C8W2
22	69	10.7	454	13	090ZEB
23	69	10.7	527	10	09SKW4
24	69	10.7	1081	6	09BE73
25	68.5	10.6	381	5	026641
26	68.5	10.6	405	13	09PS79
27	68.5	10.6	422	11	0922X0
28	68.5	10.6	427	6	095MH5
29	68.5	10.6	1006	16	P73012
30	68	10.6	189	14	0991U7
31	68	10.6	283	10	043508
32	68	10.6	291	13	091AK3
33	68	10.6	539	10	09XHV3
34	67.5	10.5	453	13	09FTN2
35	67.5	10.5	1069	5	09BPS2
36	67	10.4	376	10	09LXW3
37	67	10.4	454	13	08GGE2
38	67	10.4	743	12	083111
39	67	10.4	1062	11	088845
40	67	10.4	1062	11	089019
41	66.5	10.3	103	17	09YC22
42	66.5	10.3	419	11	062681
43	66.5	10.3	422	13	091430
44	66.5	10.3	2571	11	08R4V4
45	66.5	10.3	3301	11	091Z10

ALIGNMENTS

09fWv9 Oryza sativ
09hVU8 pseudomonas
09hBS0 homo sapien
074213 aspergillus
09c8W2 arabidopsis
090ZEB oncorhynch
09SKW4 arabidopsis
09BE73 macaca fasc
026641 strongyloce
09PS79 xenopus. xc
0922X0 mus musculu
095MH5 sagittus oe
P73012 synechocyst
0991U7 uncultured
043508 lycopersico
091AK3 xenopus lae
09XHV3 Oryza sativ
09FTN2 brachydanio
09BPS2 bombyx mori
09LXW3 arabidopsis
08GGE2 salmo salar
083111 mouse adeno
088845 mus musculu
089019 mus musculu
09YC22 aeropyrum p
062681 rattus norv
091430 brachydanio
08R4V4 mus musculu
091Z10 mus musculu

RESULT 1
08R3U6 PRELIMINARY; PRT; 128 AA.
AC 08R3U6;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to hypothetical gene LOC125944.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024561; AAH24561.1; -
SQ SEQUENCE 128 AA; 14043 MW; CC6E7DA04D3D7DF CRC64;

Query Match 50.0%; Score 322; DB 11; Length 128;
Best Local Similarity 70.9%; Pred. No. 2.1e-29;
Matches 61; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

QY 1 MKVLISLLPLPLMLSMVSSSLNPGVARGHNRGASRRVLOEGQEECKDMFIRAP 60
DB 1 MKLIASFLLPLVLMISWFSPPNPGVARGHDIAPRRWLLEGGQEECKDMFIRAP 60
QY 61 RRFMTVSGLPKRCPCDHFKNYK 86
DB 61 KKKATAVLGPFRKCCPDHYKREKK 86
RESULT 2
AC 062022 PRELIMINARY; PRT; 477 AA.
ID 062022;
DT 01-ANG-1998 (TREMBLrel. 07, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

```

DE ZK131.11 protein.
GN ZK131.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Percy C.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z93373; CAB07552.1; -.
DR EMBL; Z83245; CAB07552.1; JOINED.
DR EMBL; Z83245; CAB05840.1; -.
DR EMBL; Z93373; CAB05840.1; JOINED.
SO SEQUENCE 477 AA; 53385 MW; 86F5257F9EB6DD90 CRC64;

Query Match 12.0%; Score 77.5; DB 5; Length 477;
Best Local Similarity 34.7%; Pred. No. 1.2;
Matches 25; Conservative 6; Mismatches 32; Indels 9; Gaps 3;

OY 46 GGOECCDMFLRAPP--KFMVSGLPKKO--CPCDHFKNVKKTRHQHRRKPKHS 100
DB 290 GGPANESNTTFIRATKRFKKITTSATLPKKNSTSMDFQKSVFPEDTAHNNRMDG 349
OY 101 RACQOFLKQCOL 112
DB 350 R---ILKKAPL 357

RESULT 3
OY 09W745 PRELIMINARY; PRT; 411 AA.
AC G9W745;
DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
DT 01-NOV-1999 (TREMBLrel. 21, last annotation update)
DE COUP transcription factor 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Langlois M.C., Vanacker J.M., Holland N.D., Escriva H., Queva C.,
RT "Amphicoupe-tr, a nuclear orphan receptor of the lancelet Branchiostoma
RT floridae, is implicated in retinoic acid signaling pathways."
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AF157558; AAD4224.1; -.
DR HSSP; P19793; 2NLT.
DR InterPro; IPR000536; Hormone_rec_1lg.
DR InterPro; IPR001628; Znf_Cassteroid.
DR Pfam; PF00104; hormone_rec_1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STEROIDFINGER.
DR PRODOM; PD000035; Znf_Cassteroid; 1.
DR SMART; SM00430; HOLY. 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SO SEQUENCE 411 AA; 45467 MW; 950E285B7FC336FB CRC64;

Query Match 11.9%; Score 76.5; DB 13; Length 411;
Best Local Similarity 26.0%; Pred. No. 1.4;
Matches 26; Conservative 16; Mismatches 25; Indels 33; Gaps 6;

OY 26 PGVA--RGRHDSGASRRRLQ-----EGGQECF--CQDMFLRARRKFMVSG 70
DB 53 PGAAADGGQGGGSOQHIECVVCGDKSSGKHGYGTFEGCGSKFPRKSVARN-LVYTCR 111

```

```

OY 71 PKKQPCDHFKNVKKTRHQHRRKPKHSRACQOFLKQC 110
DB 112 ANNCPTID-----QHHRNCCQYCR-----LKRC 134

RESULT 4
OY 088278 PRELIMINARY; PRT; 3313 AA.
AC 088278;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE MEGF2.
GN MEGF2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
RX MEDLINE-98360089; PubMed-9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RT like motifs by motif-trap screening."
RL Genomics 51:27-34(1998).
CC -1- SIMILARITY: CONTAINS 9 CADHERIN DOMAINS.
DR EMBL; AB011528; BAA32459.1; -.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormu_receptor.
DR InterPro; IPR002049; laminin_BGF.
DR InterPro; IPR001791; laminin_G.
DR InterPro; IPR000203; PKD_cytrich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF00028; cadherin; 9.
DR Pfam; PF00008; EGF; 6.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00054; laminin_G; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 8.
DR SMART; SM00180; EGF_Lam; 1.
DR SMART; SM00001; EGF_Like; 6.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00282; Lamg; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00232; CADHERIN_1; 6.
DR PROSITE; PS0268; CADHERIN_2; 9.
DR PROSITE; PS00022; EGF_1; UNKNOWN_6.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS02227; G_PROTEIN_RECPT_P2_3; 1.
DR PROSITE; PS0261; G_PROTEIN_RECPT_P2_4; 1.
DR PROSITE; PS01248; LAMININ_TYRE_EGF; UNKNOWN_1.
DR Calcium-binding; Cell adhesion; EGF-like domain; Glycoprotein.
KW SEQUENCE 3313 AA; 359348 MW; B1DA09517288764 CRC64;

Query Match 11.3%; Score 72.5; DB 11; Length 3313;
Best Local Similarity 27.6%; Pred. No. 35;
Matches 35; Conservative 9; Mismatches 40; Indels 43; Gaps 7;

OY 24 LNFVGARHGDRGQASRRWLQEG--GQECCKDMFL-----RAPRRFMVSG 71
DB 1978 LNF-----CQNGSGRHLQGGPHGYTCDCAAGYFGQCHHRMDQCPGRGMWSPTGCP 2030
OY 72 KKQPCDHFKNVKKTRHQHRRKPKHSRACQOFLKQC 110

```

```

Db 2031 ---CNCVDVHGFDPNCKTSGCHCEKHFHPRGSDSCLEPCDYPWGSTRSCAPHSQGC 2087,
OY 111 QLRSPAL 117
Db 2088 PCRRGAL 2094

RESULT 5
082328
ID 082328 PRELIMINARY; PRT; 108 AA.
AC 082328;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE GAST1/GASA-like protein.
GN AT2G14900.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman J.C., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005396; AAC61287.1; -.
DR InterPro: IPR003854; GASA.
DR Pfam: PF02704; GASA; 1.
SQ SEQUENCE 108 AA; 11381 MW; 93206EAE3BB95DF5 CRC64;

Query Match 11.1%; Score 71.5; DB 10; Length 108;
Best Local Similarity 25.9%; Pred. No. 1.3;
Matches 29; Conservative 18; Mismatches 38; Indels 27; Gaps 7;

OY 1 MKVLSLLLL-LPLIMSKVSSSLNPGVARG-----HRDRGQASRRMLQEGGEGE-- 51
Db 1 MKTIVSILVLAASLLISLASATISDAPSGAVAPAPSKDPALEKMC---GQCEGR 57
OY 52 CKDWFLAPRRKFTVSGLPKKQCPDHFEGNVKKTTHORHHRKPKKHSRA 103
Db 58 CKE---AGMKDRCLKCYCKCKDQC-----VPSTY-----GKKHECAC 94

RESULT 6
08TEV2
ID 08TEV2 PRELIMINARY; PRT; 188 AA.
AC 08TEV2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Vascular endothelial growth factor B isoform VEGF-B167.
GN VEGFB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=96325041; PubMed=8702615;
RA Olofsson B., Pajusola K., von Euler G., Chliov D., Allitalo K.,
RA Eriksson U.;
RT "Genomic organization of the mouse and human genes for vascular
RT endothelial growth factor B (VEGF-B) and characterization of a second
RT splice isoform."
RL J. Biol. Chem. 271:19310-19317(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98032228; PubMed=9365524;
RA Joukov V., Kaipainen A., Jeltsch M., Pajusola K., Olofsson B.,
RA Kumar V., Eriksson U., Allitalo K.;
RT "Vascular endothelial growth factors VEGF-B and VEGF-C."
RL J. Cell. Physiol. 173:211-215(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Jeltsch M., Pajusola K., Chliov D., Allitalo K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF468110; AAL79000.1; -.
SQ SEQUENCE 188 AA; 21314 MW; F04654D5A3626095 CRC64;

Query Match 11.0%; Score 71; DB 4; Length 188;
Best Local Similarity 29.7%; Pred. No. 2.7;
Matches 22; Conservative 13; Mismatches 15; Indels 24; Gaps 6;

OY 43 LOEGGCECKDFLAPRRKFTVSGLPKKQCPDHFEGNVKKTTHORHHRKPKKHSR 101
Db 116 LEEHSQ-CECR-----PKKDSNVKPSRPLCP-----KCTGNHQRD--PR 155
OY 102 ACOQFLKQCLRSF 115
Db 156 TCR---RCRRRSF 166

RESULT 7
001368
ID 001368 PRELIMINARY; PRT; 3190 AA.
AC 001368;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE CREB-binding protein homolog.
GN NEJ OR CG15319.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97263578; PubMed=9109493;
RA Akimaru H., Chen Y., Dai P., Hou D.X., Nonaka M., Smolik S.M.,
RA Armstrong S., Goodman R.H., Ishii S.;
RT "Drosophila CBP is a co-activator of cubitus interruptus in hedgehog
RT signaling."
RL Nature 386:735-738(1997).
DR EMBL: U88570; AAB53050.1; -.
DR TRANSFAC: T03236;
DR FlyBase: Fggn0015624; nej.
DR InterPro: IPR001487; Bromodomain.
DR InterPro: IPR003101; KIX.
DR InterPro: IPR000197; TAZ_finger.
DR InterPro: IPR000433; ZnF_ZZ.
DR Pfam: PF00439; bromodomain; 1.
DR Pfam: PF02172; KIX; 1.
DR Pfam: PF02135; zf-TAZ; 2.
DR Pfam: PF00569; zf; 1.
DR PRINTS: PR00503; BROMODOMAIN.
DR SMART: SM00297; BROMO; 1.
DR SMART: SM00291; ZnF_ZZ; 1.
DR PROSITE: PS00633; BROMODOMAIN_1; UNKNOWN_1.
DR PROSITE: PS50014; BROMODOMAIN_2; 1.

```

DR PROSITE: PS01357; ZF-ZZ.1; UNKNOWN.1
 SQ SEQUENCE 3190 AA; 331879 MW; E53526678BC055A8 CRC64;
 Query Match 11.0%; Score 71; DB 5; Length 3190;
 Best Local Similarity 26.7%; Pred. No. 50;
 Matches 23; Conservative 11; Mismatches 30; Indels 22; Gaps 3;

QY 25 NPGVARGHGRGQASRRWLDGSGDECKDWFRLARRKFTVSGLPKKQCCDFHFGNV 84
 DB 2390 NPGQA-----RKQSIQKCIOSLAHACCCRDANCRLP-----SCQMKLIV 2429
 QY 85 KTRHQRHRRKPNKHSRACQOFLKQC 110
 DB 2430 QHTKCK--RKPNGCCPICQKQLALC 2453

RESULT 8
 Q9W321 PRELIMINARY; PRT: 3275 AA.
 AC Q9W321;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CG15319 protein.
 GN NEI OR CG15319.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=2019600; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Abritl J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Beus P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson R., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostel R., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Keichum K.A.,
 RA Jaimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svendsen R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 DB EMBL; AE003448; AAF46516.1; -.

DR FlyBase: FBgn0015624; nef.
 DR InterPro: IPR001487; Bromodomain.
 DR InterPro: IPR003101; KIX.
 DR InterPro: IPR000197; TRAFinger.
 DR InterPro: IPR000433; Znf_ZZ.
 DR Pfam: PF00439; bromodomain; 1.
 DR Pfam: PF02172; KIX; 1.
 DR Pfam: PF02135; ZF-TAZ; 2.
 DR Pfam: PF00569; ZZ; 1.
 DR PRINTS: PR00503; BROMODOMAIN.
 DR SMART: SM00287; BROMO; 1.
 DR SMART: SM00281; ZNF_ZZ; 1.
 DR PROSITE: PS00633; BROMODOMAIN.1; 1.
 DR PROSITE: PS50014; BROMODOMAIN.2; 1.
 DR PROSITE: PS01357; ZF-ZZ.1; UNKNOWN.1.
 SQ SEQUENCE 3275 AA; 340672 MW; E5944C3BFEC0E7AA CRC64;

Query Match 11.0%; Score 71; DB 5; Length 3275;
 Best Local Similarity 26.7%; Pred. No. 51;
 Matches 23; Conservative 11; Mismatches 30; Indels 22; Gaps 3;

QY 25 NPGVARGHGRGQASRRWLDGSGDECKDWFRLARRKFTVSGLPKKQCCDFHFGNV 84
 DB 2389 NPGQA-----RKQSIQKCIOSLAHACCCRDANCRLP-----SCQMKLIV 2428
 QY 85 KTRHQRHRRKPNKHSRACQOFLKQC 110
 DB 2429 QHTKCK--RKPNGCCPICQKQLALC 2452

RESULT 9
 Q9SQH1 PRELIMINARY; PRT: 160 AA.
 AC Q9SQH1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Allergen.
 GN ARA H 7.
 OS Atrachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Atrachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VIRGINIA;
 RX MEDLINE=99406463; PubMed=10474031;
 RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;
 RT "Selective cloning of peanut allergens, including profilin and 2S
 RT albumin, by phase display technology.";
 RL Int. Arch. Allergy Immunol. 119:265-274 (1999).
 DR EMBL: AF091737; AAD56719.1; -.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; TRY/amy1_inhbr.
 DR Pfam: PF00234; try_alpha_amy1; 1.
 DR SMART: SM00499; AAI; 1.
 SQ SEQUENCE 160 AA; 18417 MW; 9F9E4CEB68080DAC CRC64;

Query Match 10.9%; Score 70.5; DB 10; Length 160;
 Best Local Similarity 22.1%; Pred. No. 2.6;
 Matches 33; Conservative 19; Mismatches 46; Indels 51; Gaps 6;

QY 1 MKVLISLLLLPLIMSVSSSLNPGVARGHGRGQASRRWLDGSGDECKDWFRLAP 60
 DB 1 MAAKLSIIVALLGALIVASATRMP-----DRSGSGSRWDAPSRGDCQORLQIRAN 53
 QY 61 RRRKFTVSGLPKKQCCDFHFGNVKTRHQRHRRKPNKHSRAC 103
 DB 54 LR-----PCEHHRRRVRVEQBOEDYRSTRSGSGRGRPGESDENQEDRCC 100
 QY 104 Q-----QFLKO-CQLRSFALP 118

Db 101 NELNRFONNORCMQALQILQWQSFVWP 129

RESULT 10

096DA3 PRELIMINARY; PRT; 1290 AA.
AC 096DA3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Hypothetical 137.8 kDa protein.
GN B11822.060.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algen V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Myakura G., Neues H.W., Mannhaupt G.;
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL EMBL; AL356834; CAD11425.1; -
KW Hypothetical protein.
SQ SEQUENCE 1290 AA; 137768 MW; 1076A879259BEB70 CRC64;

Query Match 10.9%; Score 70.5; DB 3; Length 1290;
Best Local Similarity 24.5%; Pred. No. 22;
Matches 26; Conservative 18; Mismatches 31; Indels 31; Gaps 6;

QY 33 RDGQASRRRLQEGGCECKDMFLAPRRKFTVS---GLPKQCCDFKGNVKK--T 87
Db 454 RNNKKMKKQKWKQSGSKS-----LRAPESAMPVSVYVWSPSPRAWPLPRQGHIRNGDV 508
QY 88 RHQRHRR-----KPNKHSRACQOF--LKOCQLRSFA 116
Db 509 GCHSHQNHMYNDEGYVRRLMSEKLEPH---QQFARLRKALALXA 550

RESULT 11

09BRN8 PRELIMINARY; PRT; 117 AA.
AC 09BRN8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Hypothetical 13.1 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006151; AAH06151.1; -
KW Hypothetical protein.
SQ SEQUENCE 117 AA; 13085 MW; 77A87C1DC40665DC CRC64;

Query Match 10.9%; Score 70; DB 4; Length 117;
Best Local Similarity 29.5%; Pred. No. 2.1;
Matches 26; Conservative 7; Mismatches 23; Indels 32; Gaps 5;

QY 16 LMSVSSSLNPGYAR---GHRDRGQASRRWLQEGGCECKDMFLAPRRKFTVSGIPK 72
Db 18 LLSMV-----PGPARPPGSCWDPTQCTRTWL-----LSHTPRRRW--ISGLPR 58
QY 73 KQCP-----CDHFKGNVKKTRHQ 90

Db 59 ASCRLGEPPPLPYCDQAYGEELSTRHR 86

RESULT 12

097120 PRELIMINARY; PRT; 743 AA.
AC 097120;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Retinoid acid receptor RXR.
DE Retinoid acid receptor RXR.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99143112; PubMed-9988692;
RA Freebern W.J., Osman A., Niles E.G., Christen L., LeVerde P.T.;
RT Identification of a cDNA encoding a retinoid X receptor homologue
RT from Schistosoma mansoni. Evidence for a role in female-specific gene
RT expression. J. Biol. Chem. 274:4577-4585(1999).
RL J. Biol. Chem. 274:4577-4585(1999).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF094759; AAD16119.1; -
DR HSSP; P19793; 1RXR.
DR InterPro; IPR000536; Hormone_rec_119.
DR InterPro; IPR001723; Steroid_receptor.
DR InterPro; IPR001628; ZnF_C4steroid.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; ZnF_C4steroid; 1.
DR SMART; SM00430; HOD1; 1.
DR SMART; SM00399; ZnF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 743 AA; 82558 MW; 4CA9D287270B64DA CRC64;

Query Match 10.9%; Score 70; DB 5; Length 743;
Best Local Similarity 28.4%; Pred. No. 14;
Matches 29; Conservative 15; Mismatches 28; Indels 30; Gaps 7;

QY 18 SWSVSSSLNP-GVARGHRDQASRRWLQEGGCECKDMFLAPRRKFTV---SGLPK 72
Db 261 SVYTTNLPICVIGCDKASGK-----HYGVISCEGCKGFFKRTVRRQLVYVCRSG--- 311
QY 73 KQCP-----CDH-----FKGNVKKTRHQHHRKPKKH 99
Db 312 -QCPVDRRKRTQCHRFEDQCLAKMKKEAVQERHROPSSN 352

RESULT 13

021852 PRELIMINARY; PRT; 1095 AA.
AC 021852;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE R09A8.1 protein.
GN R09A8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkison J.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; 268009; CAA92003.1;
 SQ SEQUENCE 1095 AA; 124431 MW; 33DD93BD906752C CRC64;

Query Match 10.9%; Score 70; DB 5; Length 1095;
 Best Local Similarity 31.6%; Pred. No. 21;
 Matches 24; Conservative 10; Mismatches 32; Indels 10; Gaps 3;

OY 20 VSSSLNPGVARGHRDQASRRMLQBGQCECKDFLAPRRKFMVSGLPKKQCPD 79
 DB 305 VOSSLMVGNGRSTDDRVNA-QRPASTGFEKRCRMDQLVEKKO-----KKEVNSDH 355
 OY 80 EKGNVKTRHQRHRRK 95
 DB 356 KKAN-RITSHLEHNSR 370

RESULT 14

O9TV02 PRELIMINARY; PRT; 1664 AA.

ID O9TV02;
 AC 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Y64G10A.7 protein.
 GN Y64G10A.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermata; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Morimoto B.J.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; 268009; CAA92003.1;
 SQ SEQUENCE FROM N.A.

RA Ainscough R.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL117206; CAB60454.1; JOINED.
 DR EMBL; AL110498; CAB57911.1; JOINED.
 DR EMBL; AL110498; CAB57911.1; JOINED.
 DR EMBL; AL117206; CAB57911.1; JOINED.
 DR HSSP; P00736; IAPQ.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF Ca.
 DR InterPro; IPR002049; Laminin_EGF.
 DR Pfam; PF00008; EGF_25.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR SMART; SM00179; EGF_CA; 4.
 DR SMART; SM00001; EGF_like; 18.
 DR PROSITE; PS00010; ASX_HYDROXYL; 4.
 DR PROSITE; PS00022; EGF_2; UNKNOWN_22.
 DR PROSITE; PS01186; EGF_2; 24.
 DR PROSITE; PS01187; EGF_CA; 3.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 KW Calcium-binding; EGF-like domain; A69F093B4C705832 CRC64;
 SQ SEQUENCE 1664 AA; 179279 MW; A69F093B4C705832 CRC64;

Query Match 10.9%; Score 70; DB 5; Length 1664;
 Best Local Similarity 30.0%; Pred. No. 33;

Matches 21; Conservative 10; Mismatches 35; Indels 4; Gaps 3;

OY 46 GGOBCECKMWF-LRAPRRKFMVSGLPKKQCPDHFHKGKVKTRH--ORHHRKPKHNSRA 102
 DB 273 GGAQCCQCPGFHLSYDRSCVDIDECANN--GCHFCENVKGTTRCKCKEYQLGRDGR 331
 OY 103 CQOFLKQCOL 112
 DB 332 CEMLGSCQV 341

RESULT 15

O85299 PRELIMINARY; PRT; 221 AA.

ID O85299;
 AC 085299;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
 DE Homologue of retroviral pseudoprotease.
 OS Orf virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Parapoxvirus.
 OX NCBI_TaxID=10258;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90266454; PubMed-2129563;
 RA Fraser M., Hill D.F., Mercer A.A., Robinson A.J.;
 RT "Sequence analysis of the inverted terminal repetition in the genome
 of the parapoxvirus, orf virus."
 RL Virology 176:379-389(1990).
 DR EMBL; M30023; AAA46789.1;
 SQ SEQUENCE 221 AA; 25266 MW; 25659ABD386C2317 CRC64;

Query Match 10.8%; Score 69.5; DB 12; Length 221;
 Best Local Similarity 29.2%; Pred. No. 4.7;
 Matches 33; Conservative 8; Mismatches 37; Indels 35; Gaps 6;

OY 26 PGVARGHRDQASRRMLQBGQCECKDFLAPRRKFMVSGLPKKQCPD 79
 DB 83 PAARAGARRRCCARRARGGWP--RRWRPARRG--SGPARAPAAALADPAD 135
 OY 77 CDHFKNVKKTR-----HOR--HHRKPKHNSRAQOFLKQCOLS 114
 DB 136 ---RSKYKTRLAVETLPPQPRTPHRLPPARRQHSQAACIPRARGCSARS 184

Search completed: April 16, 2003, 12:16:13
 Job time : 36 secs